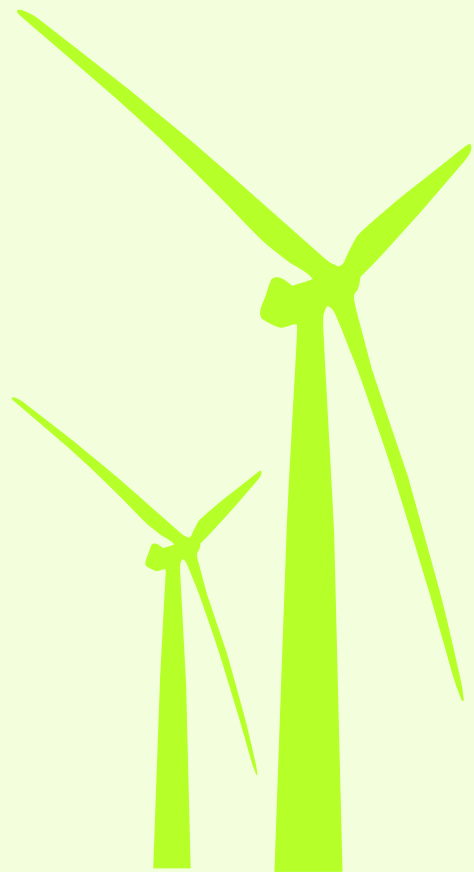


SQL Project

Pollution and Renewable
Energy Sources

Elisa Atzori



Introduction

In this analysis, we will explore themes of vital importance for our planet: pollution and renewable energy sources. Through a data analysis using SQL, we will delve into the dynamics that define the character of each nation.

The main objective is to examine the correlation between pollution, measured through CO2 emissions, and the adoption of renewable energy sources. To gain a comprehensive understanding, we will also examine a range of demographic, economic, and geographic variables. This will allow us to comprehend how nations interact with their environment and with each other, providing a broader picture of the challenges and opportunities related to environmental sustainability.

SQL SCRIPT: <https://github.com/eli-atz/SQL-Project/blob/main/Script%20SQL%20-%20Pollution%20and%20Renewable%20Energy%20Sources.sql>



country_data

```
CREATE TABLE IF NOT EXISTS country_data (  
  Country VARCHAR(255),  
  Abbreviation VARCHAR(2),  
  Land_Area_Km2 INTEGER,  
  Armed_Forces_Size INTEGER,  
  Capital_Major_City VARCHAR(255),  
  Currency_Code VARCHAR(5),  
  GDP INTEGER,  
  Largest_City VARCHAR(255),  
  Official_Language VARCHAR(255),  
  Population INTEGER  
);
```

demographic_data

```
CREATE TABLE demographic_data (  
  Country VARCHAR(255),  
  Density_Per_Km2 FLOAT,  
  Birth_Rate FLOAT,  
  Fertility_Rate FLOAT,  
  Gross_Primary_Education_Enrollment FLOAT,  
  Gross_Tertiary_Education_Enrollment FLOAT,  
  Life_Expectancy FLOAT,  
  Minimum_Wage FLOAT,  
  Population INTEGER,  
  Population_Labor_Force_Participation FLOAT,  
  Unemployment_Rate FLOAT,  
  Urban_Population INTEGER  
);
```

Tables creation

The first step of the analysis process involves **creating the necessary tables**. In this phase, columns and their respective data types are carefully chosen to ensure a consistent data structure.

Once the tables are created, the next step is to import the CSV files, allowing for a detailed exploration of the datasets through the SQL language.

Tables:

- country_data
- demographic_data
- economic_data
- energy_data

economic_data

```
CREATE TABLE IF NOT EXISTS economic_data (  
  Country VARCHAR(255),  
  Abbreviation VARCHAR(2),  
  Agricultural_Land_Percentage FLOAT,  
  Co2_Emissions INTEGER,  
  Currency_Code VARCHAR(5),  
  Forested_Area_Percentage FLOAT,  
  Gdp BIGINT,  
  Gasoline_Price_Dollars FLOAT,  
  Minimum_Wage FLOAT,  
  Population INTEGER,  
  Population_Labor_Force_Participation_Percentage FLOAT,  
  Tax_Revenue_Percentage FLOAT,  
  Total_Tax_Rate FLOAT,  
  Unemployment_Rate FLOAT  
);
```

energy_data

```
CREATE TABLE energy_data (  
  entity VARCHAR(255),  
  year INTEGER,  
  electricity_from_fossil_fuels_TWh FLOAT,  
  electricity_from_nuclear_TWh FLOAT,  
  electricity_from_renewables_TWh FLOAT,  
  renewables_percent_equivalent_primary_energy FLOAT  
);
```

The two datasets under consideration are particularly extensive; therefore, I decided to eliminate some redundant columns for the type of analysis I intend to conduct, in order to streamline data management.

Data Cleaning.



Column name simplification

I decided to rename some columns to simplify future queries.

```
ALTER TABLE demographic_data
RENAME COLUMN population_labor_force_participation TO population_labor_force_perc;

ALTER TABLE demographic_data
RENAME COLUMN gross_primary_education_enrollment TO primary_education_enroll;

ALTER TABLE demographic_data
RENAME COLUMN gross_tertiary_education_enrollment TO tertiary_education_enroll;

ALTER TABLE economic_data
RENAME COLUMN population_labor_force_participation_percentage TO population_labor_force_perc;

ALTER TABLE energy_data
RENAME COLUMN renewables_percent_equivalent_primary_energy TO renewables_percent;
```

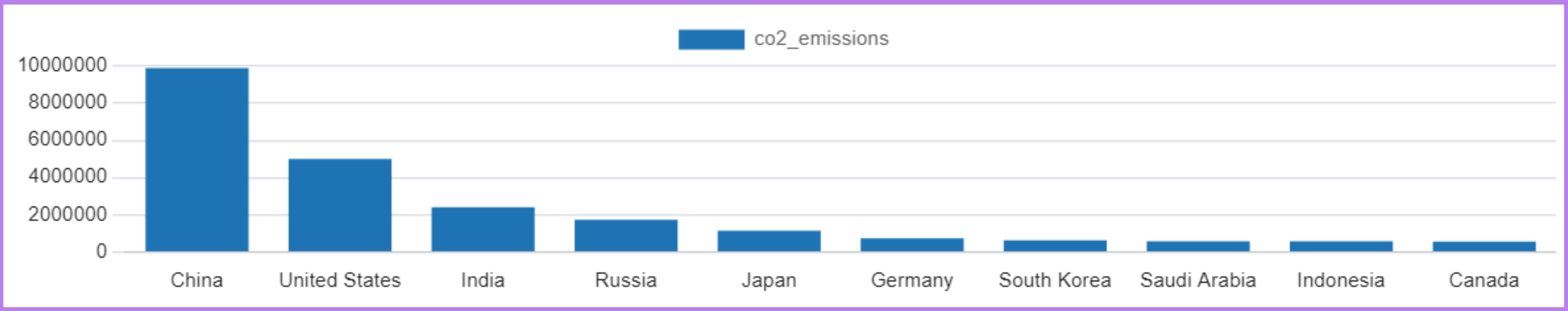
Data Analysis



Top 10 countries with highest CO2 emissions

```
SELECT country, co2_emissions
FROM economic_data
WHERE co2_emissions IS NOT NULL
ORDER BY co2_emissions DESC
LIMIT 10;
```

	country character varying (255)	co2_emissions integer
1	China	9893038
2	United States	5006302
3	India	2407672
4	Russia	1732027
5	Japan	1135886
6	Germany	727973
7	South Korea	620302
8	Saudi Arabia	563449
9	Indonesia	563325
10	Canada	544894



Nations can have high carbon dioxide (CO₂) emissions for several reasons:

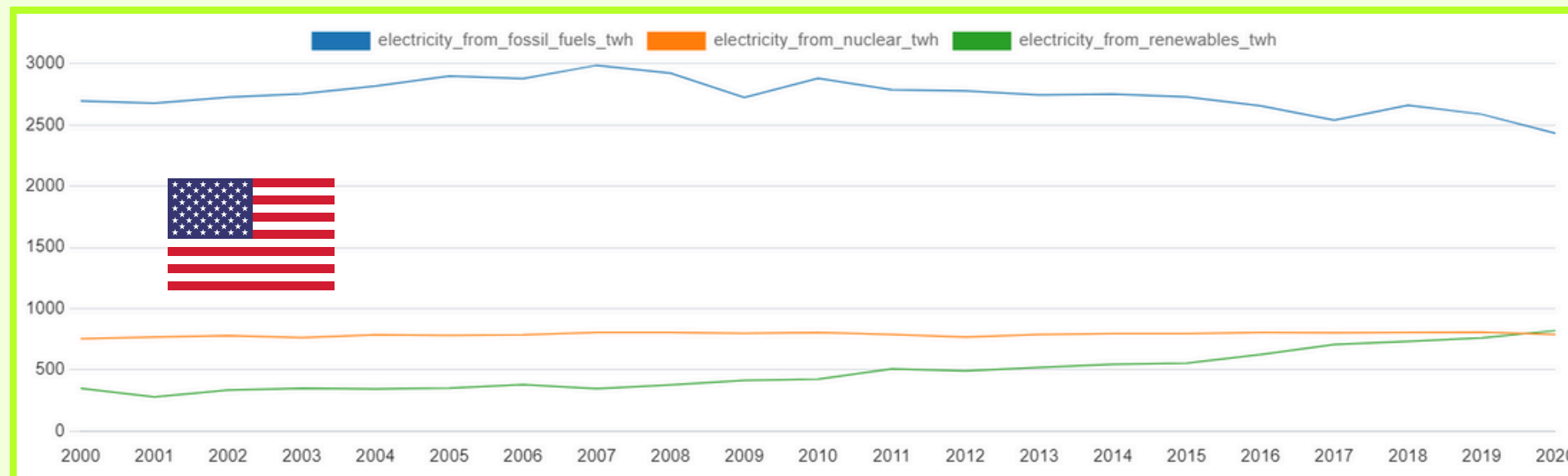
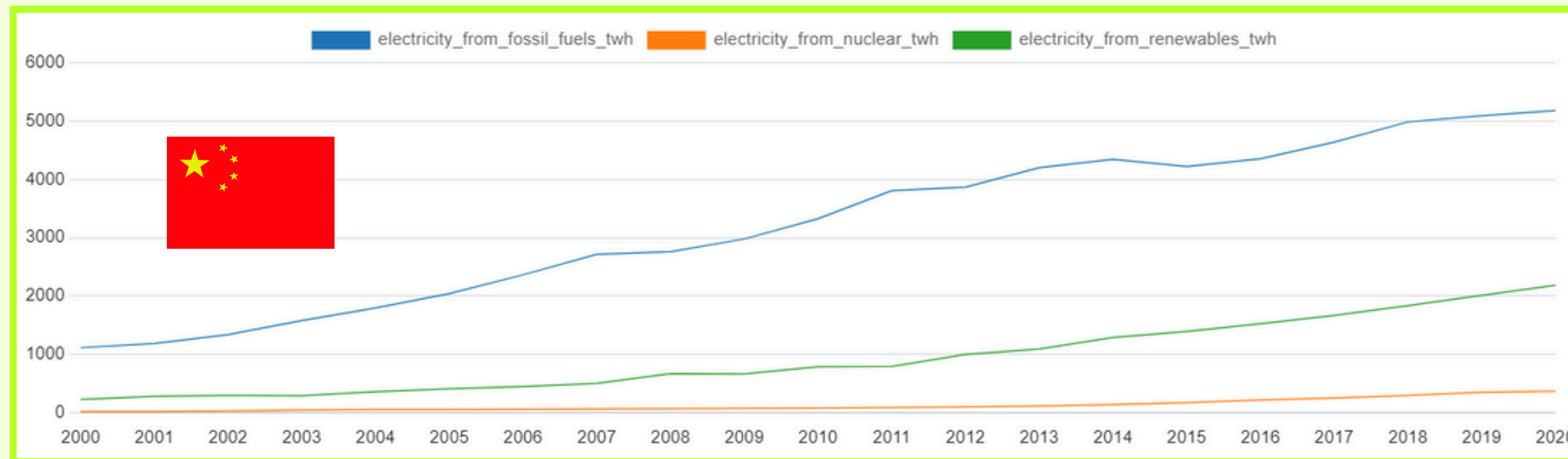
- **Population size and economic development:** a large population and/or a highly developed economy can lead to an increase in energy demand, industrial production, and resource consumption.
- **Industrial structure:** broad and diversified industrial sectors, including manufacturing, natural resource extraction, and heavy industry.
- **Dependency on fossil fuels:** many of these countries heavily rely on fossil fuels such as coal, oil, and natural gas for energy production and the functioning of their industries.
- **Mobility and transportation:** high dependence on motor vehicles and widespread transportation significantly contribute to greenhouse gas emissions in these countries.
- **Technological development and lack of energy efficiency:** in some cases, rapid industrial development may not be accompanied by advanced technologies and efficient energy practices, leading to a less sustainable use of energy resources and increased CO₂ emissions.



Let's take **China** and the **United States** as examples:

```
SELECT entity, year, electricity_from_fossil_fuels_twh, electricity_from_nuclear_twh, electricity_from_renewables_twh
FROM energy_data
WHERE entity = 'China'
ORDER BY year ASC;
```

United States' query is in the script!



In the last two decades, **China** has experienced a consistent growth in the use of fossil fuels for electricity generation, with values significantly higher than the utilization of renewable and nuclear energy. In contrast, the **United States** has maintained a stable consumption of fossil fuels over the years, with the latter consistently prevailing.

Over time, however, there has been an increase in the use of renewable sources. This shift has been driven by more stringent environmental policies, a transition towards renewable energy sources, international commitment to emission reduction, and advancements in clean technologies.

But unfortunately:

IL RAPPORTO IEA

Nel 2021, le emissioni di CO₂ sono aumentate del 6%: colpa di carbone e ripresa post-Covid

di Valeria Sforzini | 15 mar 2022

I have identified the **10 countries with the most populous populations**, among which, **7** also rank in the top 10 for the highest CO2 emissions. This correlation underscores the connection between a sizable population and the increase in energy demand, industrial production, and resource consumption.

Furthermore, it is noteworthy that in these states, **a significant portion of the population resides in urban areas, where there is a substantial reliance on motor vehicles** and transportation, contributing significantly to greenhouse gas emissions in such contexts.

Countries like China, the United States, and Russia, also, due to **their vast territories, require an extensive transportation network** that contributes significantly to the use of fuels.

```
SELECT c.country, c.land_area_km2, d.density_per_km2, d.population, d.urban_population_percentage
FROM country_data c
JOIN demographic_data d ON c.country = d.country
WHERE d.population IS NOT NULL AND d.urban_population_percentage IS NOT NULL
ORDER BY d.population DESC
LIMIT 10;
```

	country character varying (255) 🔒	land_area_km2 integer 🔒	density_per_km2 integer 🔒	population integer 🔒	urban_population_percentage double precision 🔒
1	China	9596960	153	1397715000	842933962
2	India	3287263	464	1366417754	471031528
3	United States	9833517	36	328239523	270663028
4	Indonesia	1904569	151	270203917	151509724
5	Pakistan	796095	287	216565318	79927762
6	Brazil	8515770	25	212559417	183241641
7	Nigeria	923768	226	200963599	102806948
8	Bangladesh	148460	1265	167310838	60987417
9	Russia	17098240	9	144373535	107683889
10	Japan	377944	347	126226568	115782416

Is GDP correlated with CO2 emissions?

```
SELECT country, gdp
FROM economic_data
WHERE gdp IS NOT NULL
ORDER BY gdp DESC
LIMIT 10;
```

	country character varying (255)	gdp bigint
1	Uruguay	214277000000000
2	China	199100000000000
3	Japan	5081769542380
4	Germany	3845630030824
5	United States	2827113184696
6	France	2715518274227
7	India	26110000000000
8	South Sudan	20290000000000
9	Italy	2001244392042
10	Brazil	1839758040766

	country character varying (255)	co2_emissions integer
1	China	9893038
2	United States	5006302
3	India	2407672
4	Russia	1732027
5	Japan	1135886
6	Germany	727973
7	South Korea	620302
8	Saudi Arabia	563449
9	Indonesia	563325
10	Canada	544894
11	Mexico	486406
12	South Africa	476644
13	Brazil	462299
14	United Kingdom	379025
15	Australia	375908
16	Turkey	372725
17	Italy	320411
18	France	303276
19	Poland	299037
20	Thailand	283763

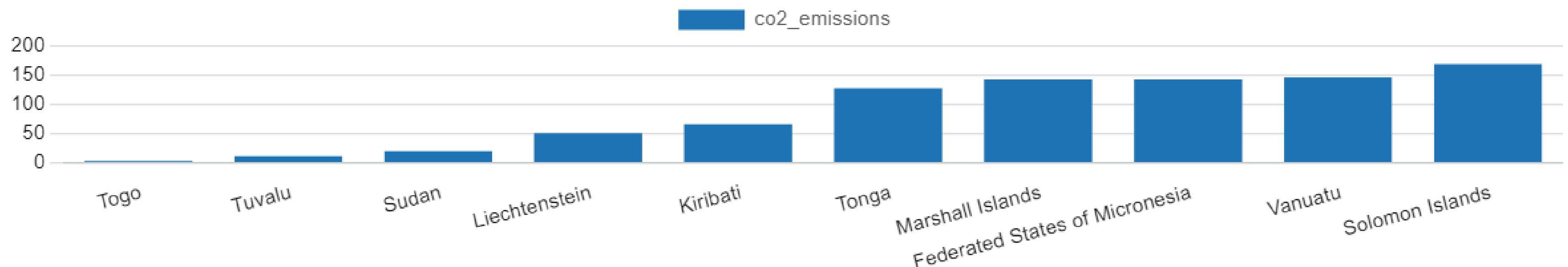
From the extractions performed, it is possible to confirm that **GDP and CO2 emissions are correlated**, as among the world's wealthiest countries, many are included in the Top 20 countries with the highest CO2 emissions.

Therefore, it is verified that wealthier countries pollute more!

Top 10 countries with lowest CO2 emissions

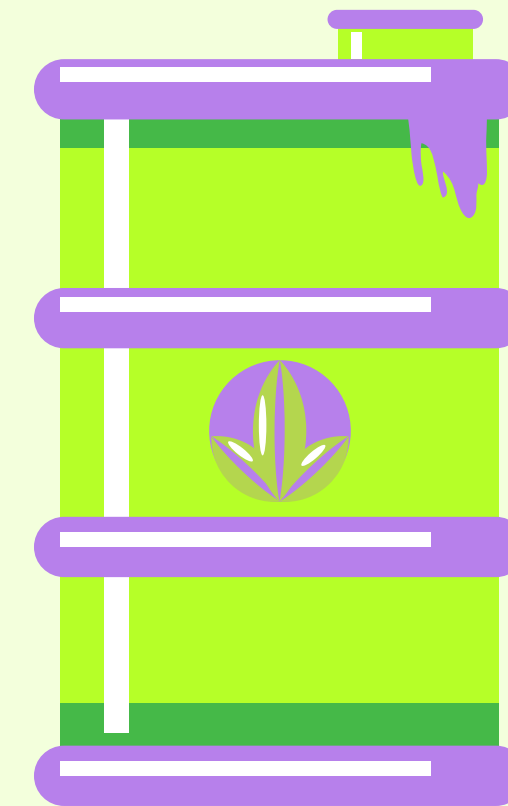
```
SELECT country, co2_emissions
FROM economic_data
WHERE co2_emissions IS NOT NULL
ORDER BY co2_emissions ASC
LIMIT 10;
```

	country character varying (255)	co2_emissions integer
1	Togo	3
2	Tuvalu	11
3	Sudan	20
4	Liechtenstein	51
5	Kiribati	66
6	Tonga	128
7	Marshall Islands	143
8	Federated States of Micronesia	143
9	Vanuatu	147
10	Solomon Islands	169



The mentioned nations generally have relatively low CO2 emissions for several reasons:

- **Population size:** Many of these nations have a relatively small population compared to larger nations, which limits total emissions.
- **Economic development:** Some of these nations may have less developed economies or economies based on sectors with lower carbon emissions, such as the service sector compared to heavy manufacturing.
- **Forested areas:** the presence of extensive forested areas in some of these nations can further contribute to carbon sequestration, helping to mitigate overall carbon emissions.
- **Dependency on renewable sources:** Some of these nations may heavily rely on renewable energy sources, such as hydroelectric, solar, or wind energy, thereby reducing CO2 emissions associated with energy production.
- **Lifestyle and resource use:** In many of these nations, traditional lifestyles and sustainable resource use can contribute to lowering per capita carbon emissions.



```
SELECT c.country, c.land_area_km2, d.density_per_km2, d.population, e.forested_area_percentage
FROM country_data c
JOIN demographic_data d ON c.country = d.country
JOIN economic_data e ON d.population = e.population
WHERE d.population IS NOT NULL
ORDER BY d.population ASC
LIMIT 20;
```



I have identified the **20 countries with the least populous populations**, among which, **6** also rank in the top 10 for the lowest CO2 emissions. Others, on the other hand, are part of the top 20.

Furthermore, all these nations share several commonalities that confirm their low CO2 emissions:

- small land areas
- low population density
- not high populations
- some of them also have very **high percentages of forested areas**.

	country character varying (255) 🔒	land_area_km2 integer 🔒	density_per_km2 double precision 🔒	population integer 🔒	forested_area_percentage double precision 🔒
1	Vatican City	0	2003	836	[null]
2	San Marino	61	566	3386	0
3	Nauru	21	541	10084	[null]
4	Tuvalu	26	393	11646	33.3
5	Palau	459	39	18233	87.6
6	Saint Lucia	616	301	18279	33.2
7	Liechtenstein	160	238	38019	43.1
8	Monaco	2	26337	38964	[null]
9	Saint Kitts and Nevis	261	205	52823	42.3
10	Marshall Islands	181	329	58791	70.2
11	Dominica	751	96	71808	57.4
12	Andorra	468	164	77142	34
13	Antigua and Barbuda	443	223	97118	22.3
14	Djibouti	23200	43	97356	0.2
15	Seychelles	455	214	97625	88.4
16	Tonga	747	147	100209	12.5
17	Saint Vincent and the Grenadines	389	284	100455	69.2
18	Grenada	349	331	112003	50
19	Federated States of Micronesia	702	784	113815	91.9
20	Kiribati	811	147	117606	15

It's important to note that many nations are adopting measures to reduce their CO2 emissions, with increased focus on renewable energy sources, energy efficiency, and environmental sustainability. However, addressing the issue comprehensively often requires long timelines and significant efforts due to the complexity of the challenges involved.

Top 10 countries with the highest percentage of renewable energy.

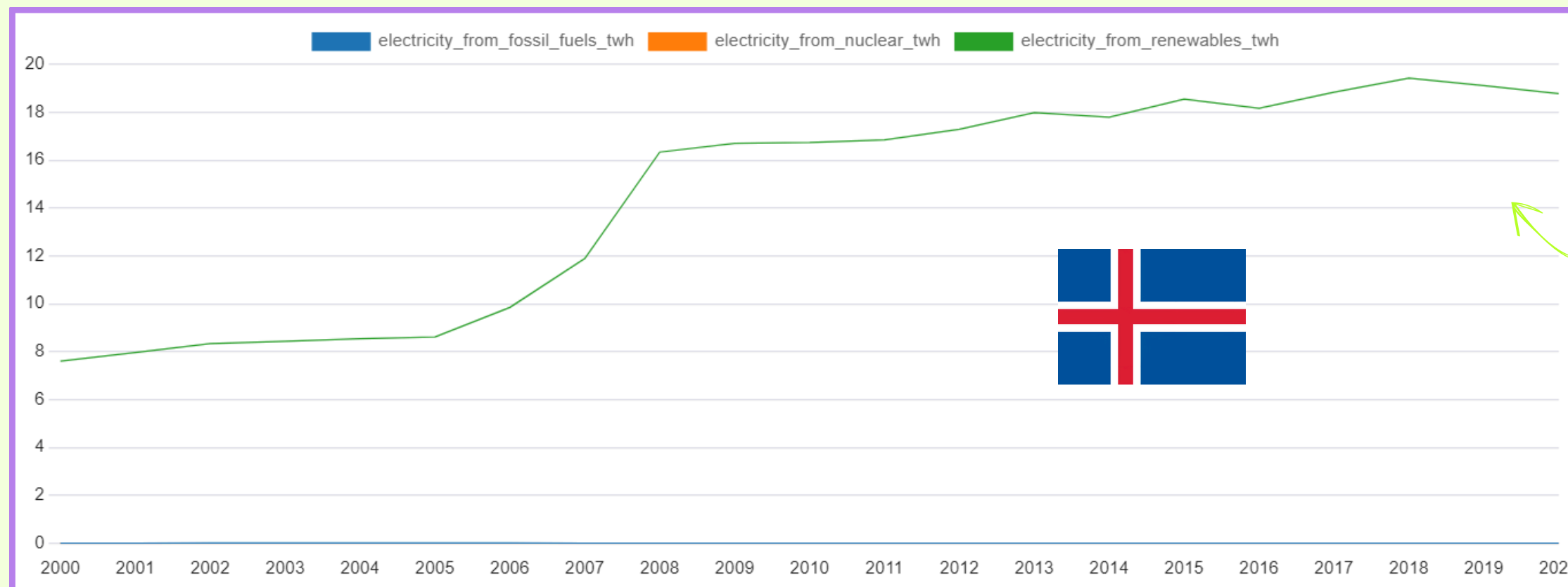
```
SELECT
    entity,
    year,
    renewables_percent
FROM energy_data
WHERE
    year = 2020
    AND renewables_percent IS NOT NULL
ORDER BY renewables_percent DESC
LIMIT 10;
```

	entity character varying (255) 🔒	year integer 🔒	renewables_percent double precision 🔒
1	Iceland ❄️	2020	86.836586
2	Norway	2020	70.96306
3	Sweden	2020	51.05897
4	Brazil	2020	49.47188
5	New Zealand	2020	39.064182
6	Austria	2020	38.26067
7	Denmark	2020	37.272453
8	Switzerland	2020	36.21057
9	Ecuador	2020	36.035774
10	Finland	2020	33.16851

It's worth noting how all the '**Nordic countries**' (Iceland, Norway, Sweden, Denmark, Finland) in Europe are present in this ranking. But why are these countries so committed to the use of renewable sources?

The Nordic countries, including Iceland, Norway, Sweden, Denmark, and Finland, have shown a strong commitment to using renewable energy sources for several reasons:

- **Abundant Natural Resources:** The region is endowed with abundant renewable energy resources, such as hydropower, wind, and geothermal energy.
- **Environmental Consciousness:** The populations in these countries are generally well-informed and concerned about environmental issues.
- **Government Policies and Support:** Governments have implemented policies and provided financial incentives to promote the use of renewable energy. This includes subsidies, tax incentives, and regulatory frameworks that encourage investments in sustainable energy solutions.



```
SELECT
  entity,
  year,
  electricity_from_fossil_fuels_twh,
  electricity_from_nuclear_twh,
  electricity_from_renewables_twh
FROM energy_data
WHERE entity = 'Iceland'
ORDER BY year ASC;
```

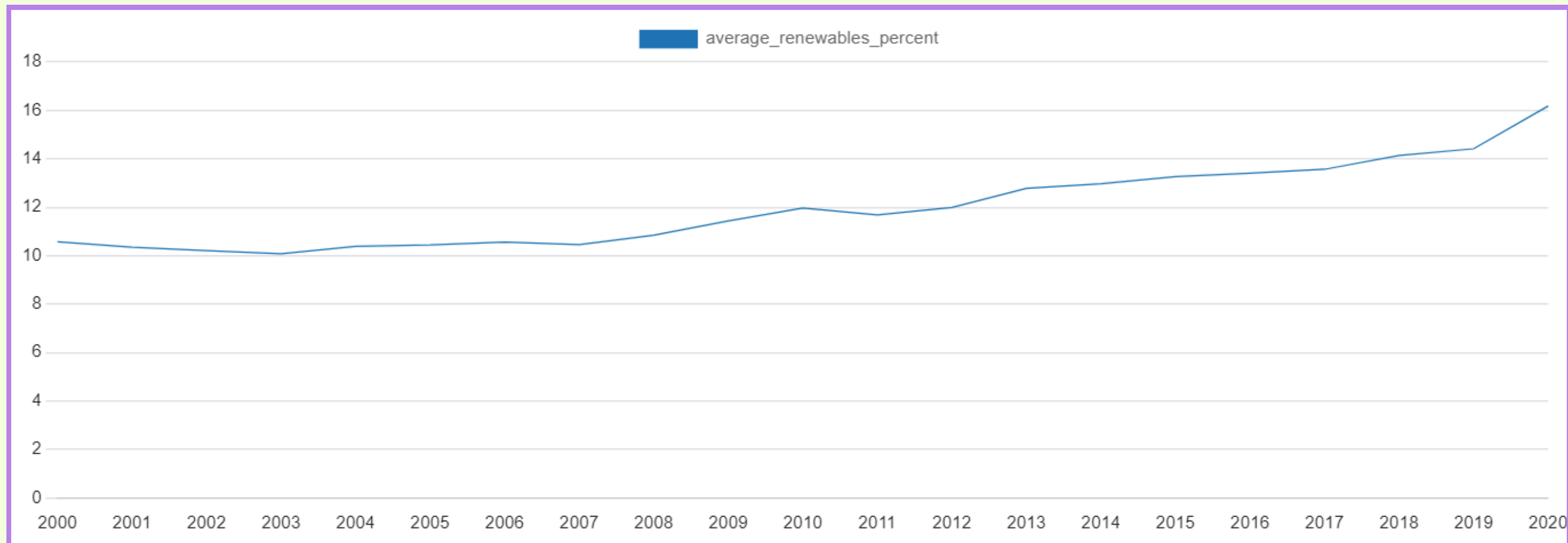
Iceland confirms the exclusive use of renewable sources for electricity supply!

Good News!!!



```
SELECT
    year,
    AVG(renewables_percent) AS average_renewables_percent
FROM
    energy_data
WHERE
    year BETWEEN 2000 AND 2020
GROUP BY
    year
ORDER BY
    year;
```

The average percentage of renewable sources globally experienced a steady positive increase during the period 2000-2020, with a growth of approximately **6%**.





Thanks

